Rotation Invariant Texture Classification with Dominant Orientation Estimation Based on Gabor Filters

Jiuwen Zhang, Xiaoyun Mai, Xiguang Wu
School of Information Science and Engineering,
Lanzhou University
Lanzhou, China, 730000

Abstract—**This paper proposes a new method to estimate the dominant orientations of textures using Gabor filters, along with its modified version to fit the multi-orientation cases. The discrete wavelet transform is used as a feature extraction tool and nearest neighbor method as a classifier. The experiments are carried out on both Brodatz database and CUReT database. The result shows that our method is effective for the classification of rotated textures and low in computational cost.**

Keywords-dominant orientation estimation; Gabor filters; discrete wavelet transform; nearest neighbor method

I. INTRODUCTION

Rotation invariant is in demand in the parttern recognition domain including texture classification, face recognition, iris recognition, fingerprint matching, etc. In recent years, many rotation invariant texture classification methods based on different viewpoints appear.

Some common methods are shown as follows.

Dominant orientation estimation method: first, estimate dominant orientations of texture images, followed by modulations. And then, extract features from those modulated images with Gabor transform [1] or wavelet transform [7] or other methods.

Feature vector modulation method in transform domain: 1.Energys of all directonal subbands from Gabor transform are calculated to form a feature vector, and the direction with the largest energy is defined as the dominant orientation. The feature vector is modulated (circularly shifted) according to that dominant orientation [5]. 2.Another way is that 1-D discrete fourier transform is carried out on the feature vector to realize rotation invariant. Because rotation of an original image is corresponding to circular shift of its featrue vector, and amplitude of a circularly shifted vector in discrete fourier transform domain is constant [2][4].

Method of Hidden Markov Models: Covariance matrix of model parameter is diagonalizable. Rotation of an original image only affects the covariance matrix, but not the diagonal matrix, which could be used to form a rotation invariant model parameter [6].

Method of combining base functions: 1.For Gabor transform, all directional base functions at each scale level are combined to form a new base function. This new base function, which is rotation invariant, is used to transform an image and extract features [3]. 2.For wavelet transform, features from HL subband and LH subband are combined to realize rotation invariance [15].

Radon transform method: 1.The Radon projection correlation distance between an original image and its rotated version is zero, but not zero between two different texture images. This distance could be used as a criterion for rotation invariant texture classification with k-nearest neighbors' classifier [9]. 2.In [11], Fourier-Mellin transform is carried out in Radon transform domain and the coefficients are modified to achieve rotation invariance.

Method of changing rotation invariance into translation invariance: 1.Radon transform is carried out on an image. Rotation of the original image is corresponding to translation along parameter θ in transform domain. Features are extracted with translation invariant wavelet transform in Radon transform domain [8][10]. 2.In the Radon transform domain, 1-D fourier transform is carried out to the distance variable t, which generates 2-D polar fourier transform of the original image. Rotation of the image is corresponding to translation in this polar fourier transform domain [14]. 3.Polar transform is carried out on an image to obtain a polar image, which means that the rotation invariance problem turns into a translation invariance problem. Then features can be extracted with many methods such as wavelet package transform [16].

Statistical method in spacial domain: Local binery pattern (LBP) is calculated from every pixel and a histogram of LBPs is built as a feature vector [12][13].

In this paper, a new dominant orientation estimation method based on Gabor filters is proposed. The outline of this paper is organized as follows: 1.A brief introduction of Gabor filters. 2.A new dominant orientation estimation method with Gabor filters and its modified version. 3.Feature extraction with discrete wavelet transform (DWT). 4.Classification with chi-square distance and nearest neighbor classifier. 5.Experiments on both Brodatz database and CUReT database.

II. GABOR FILTERS

The spacial representation and the spectral representation of a 2-D Gabor filter [3][4][5] are given as

$$g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right]. \tag{1}$$

$$G(u,v)=\exp\left\{-\frac{1}{2}\left[\frac{\left(u-W\right)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\}. \tag{2}$$

Where $\left(W,0\right)$ is the central frequency of the passband. $\sigma_u=1/2\pi\sigma_x$, $\sigma_v=1/2\pi\sigma_y$, they determine the size of the passband.

A group of filters can be generated from mother function $g(x,y)$. They are

978-1-4244-8728-8/11/$26.00 ©2011 IEEE

$$g_{m,n}(x,y) = a^{-2m} g(x', y'). \quad (3)$$

Where $x' = a^{-m}\left(x\cos\theta_n + y\sin\theta_n\right)$, $y' = a^{-m}\left(-x\sin\theta_n + y\cos\theta_n\right)$, $a > 1$, $\theta_n = n\pi / K$, $m = 0, 1, \ldots, S-1$, $n = 0, 1, \ldots, K-1$. S is the number of scales, K is the number of directions. In order to obtain the information of a certain direction, a filter that has a passband centrosymmetric to that of the filter given in (2) is generated, and these two filters are combined. Then the spacial representation of a 2-D Gabor filter becomes real [1][2].

By convolving an image with filter $g_{m,n}(x,y)$, the subband at scale m, direction n is

$$I_{m,n}\left(x,y\right) = \sum_{x'}\sum_{y'} I\left(x-x', y-y'\right) g_{m,n}\left(x', y'\right). \quad (4)$$

III. Dominant Orientation Estimation

A. Basic idea

First, If we consider the coefficients of a filter as weights, the left part of (4), $I_{m,n}\left(x,y\right)$, is a weighted sum of pixel values within neighborhood around $\left(x,y\right)$. We define such a weighted sum as an operation unit.

As shown in Fig. 1, 9 points are chosen with a same interval between them. In the neighborhood around $\left(x_i, y_j\right)$ $\left(i = 1,2,3, \quad j = 1,2,3\right)$, $I_{m,n}\left(x_i, y_j\right)$ represents the extent of alternation of texture at scale m, direction n. We define

$$n_{i,j}^{\text{dominant}} = \arg\left(\max_{n=1,\ldots,K}\left(I_{m,n}\left(x_i, y_j\right)\right)\right). \quad (5)$$

as the dominant orientation of this neighborhood, and each neighborhood has a dominant orientation of its own.

The 9 neighborhoods vote for K directions, and the direction with most votes is the dominant orientation of the texture image.

The image is modulated with this dominant orientation.

Practically, more points (neiborhoods) are choesn to make the voting more precise. In our experiments, 81 points with a same interval between them are selected. The number of directions of Gabor filters is $K = 16$, and these filters are located at scale level 2, the midfrequency.

B. Analysis and modification

For an anisotropic image (such as Fig. 2(a)), textures of all its neiborhoods are altering along the same direction. That allows our method to stand. On the other hand, rotation can not affect isotropic images (such as Fig. 2(b)).

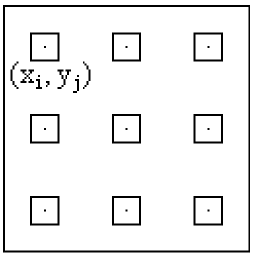


Figure 1. 9 neighborhoods on an image.

Our method demands 16×81 operation units and is lower in computational cost than traditional Radon transform method, which requires interplots and integrals. If an image has a size up to 256×256 pixels, the amount of calculation of our method is $81/65536$ as large as that of an image convolving at one scale level.

The directional minimum resolution of this method is $\pi/16$, which is sufficient for texture classificaiton.

The defination of (5) is reliable theoreticly, but not practically. Because there is more than 1 peak (dominant orientaion) among the 16 directional filtered coefficients. We modify the defination as follows: For an anisotropic image, although peaks of two neighbors may be at different directions, these two neighbors have the same coefficient distribution along 16 directions. In another way, if direction n_d has a large coefficient in neighbor around (x_{i1}, y_{j1}), it does in neighbor around (x_{i2}, y_{j2}) too. We can use the centroid of the 16 directional coefficients to define a dominant orientaion instead of the largest one. A centroid is calculated with its defination that it is the sum of positon numbers weighted by mass proportion (coefficient proportion in our case). A centroid can also be calculated with its property of having equal weights on both sides. It is noticalbe in the calculation that rotation of a neighborhood is corresponding to circular shift of directional coefficients.

IV. Feature Extraction

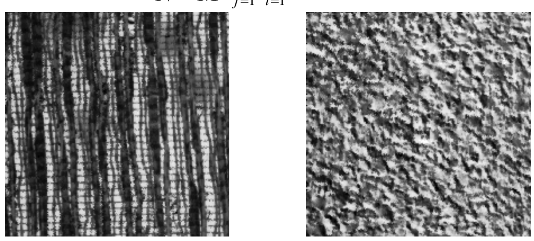
Discrete wavelet transform (DWT) is used to extract features for its high efficiency of spacial-frequency analysis, simplicity of structure, rapidity of calculation, and zero redundancy of data. DWT is deficient in multi-directional analysis abilitys for it can only decompose an image into 3 directional subbands. Furthermore, if the original image is rotated, the coefficients of all subbands change significantly. But in our experiments, we have found the dominant orientaions of images and modulated these images, and now we could extract features from these images with DWT.

2-D DWT is achieved by implementing 1-D DWT on row and column orderly. The 3 high frequency subbands are represented as LH, HL, HH.

The standard deviations of the 3 high frequency subbands are chosen as features.

$$\sigma_{s,k} = \sqrt{\frac{1}{N \times M}\sum_{j=1}^{M}\sum_{i=1}^{N}\left(I_{s,k}\left(i,j\right) - M_{s,k}\right)^2}. \quad (6)$$

$$M_{s,k} = \frac{1}{N \times M}\sum_{j=1}^{M}\sum_{i=1}^{N}\left(I_{s,k}\left(i,j\right)\right). \quad (7)$$



(a) An anisotropic image (Brodatz D51) (b) An isotropic image (Brodatz D4)

Figure 2. Anisotropic image and isotropic image.

Where $N \times M$ is the size of an image, $s = 1, 2, \ldots, S$ is the scale level, $k \in \{LH, HL, HH\}$ is the direction. $I_{s,k}(i, j)$ represents DWT coefficient of pixel (i, j) at scale level s, direction k. Then a feature vector is formed as

$$f = \left[\sigma_{1,LH}, \sigma_{1,HL}, \sigma_{1,HH}, \ldots, \sigma_{S,HH}\right]. \quad (8)$$

V. CLASSIFICATION

As a non-parameter classifier, nearest neighbor method is used to classify texture images. Its basic rule is: Find the training sample that is closest to testing sample x, and x belongs to the class of that training sample.

Chi-square distance [4][12] is used as a disimilarity measurement, and it is defined as

$$D\left(f_{train}, f_{test}\right) = \sum_{l=1}^{L} \frac{\left(\sigma_l^{train} - \sigma_l^{test}\right)^2}{\sigma_l^{train} + \sigma_l^{test}}. \quad (9)$$

Where f_{train} is the feature vector of a training sample, f_{test} is the feature vector of a testing sample, L is the dimension of a feature vector.

VI. EXPERIMENTS

Experments are carried out on Brodatz database and Columbia-Utrecht (CUReT) database to evaluate the effectiveness of our method.

A. Result on Brodatz database

We select 13 samples from Brodatz database, and they are D1, D4, D6, D19, D20, D21, D22, D52, D56, D74, D76, D102, D111. The size of these images is 640×640, and each of them is divided into 16 non-overlapping sub-images of size 128×128. Each sub-image is rotated at angle 0°, 30°, 60°, 90°, 120°, 150° to create 6 rotated sub-images. Then the total number of samples is $13 \times 16 \times 6 = 1248$. We carry out two experiments: 1. 624 images are used as training samples, and the rest 624 ones as testing samples. The proportion of training samples to testing samples is 3:3. 2. 208 images are used as training samples, and the rest 1040 ones as testing samples. The proportion of training samples to testing samples is 1:5.

In the feature extraction stage, each modulated image is decomposed with different number of scales, from1 to 5. The traditional method of Radon transform dominant orientation estimation and DWT feature extraction (similar to that is given in [7]) is tested for comparation, and is denoted as Radon-DWT (RD). Our method is denoted as Gabor-DWT (GD). The experimental result is shown in Table 1.

The dominant orientations of 1248 images are estimated with Radon transform method and Gabor filter method respectively on a same computer (same hardware configurations). The former costs us 439.63 seconds, while the latter only 77.03 seconds.

From the results, we could find that the best performances are obtained with 3 or 4 scales of DWT decomposation. More scales do not lead to better performance. This is because informations of texture images are concentrated on midfrequency, as stated in [15]. It is also observed from the table that the performance of our method is a little worse than that of Radon-DWT in terms of classification acurracy. The largest accuracys (bold) of Gabor-DWT are 2.56% and 3.94% lower than that of Radon-DWT in two experiments respectively. Because the directional minimum resolution of our method is $\pi / 16$, which is rougher compared with Radon transform method. But the amount of operations of our method is smaller, and which gives it an edge. In addition, when it comes to fewer training samples (1:5) and decomposation scales (1 or 2), our method is more accurate than Radon-DWT in classification.

B. Result on CUReT database

CUReT database is composed of images closer to real textures. These images are colorful and all images of each class are captured under varied illumination conditions and viewpoints [12][17]. Therefore, these image are rotated versions to each other and some are even scaled (along a certain direction). CUReT database is more challenging.

There are 61 classes in the CUReT database and 205 images in each class. 12 classes are selected for experiments, and they are class 4, 5, 8, 11, 15, 35, 38, 40, 43, 51, 56, 57. Among the 205 images of each class, 92 images with a viewing angle less than 60 degrees are chosen so that a sufficiently large area (128×128) could be cropped to form a sample. The total number of samples is $92 \times 12 = 1104$.

The proposed method is based on texture information and color is unused. Therefore, all images are converted to monochrome ones.

We use the following equation (from the programs with [12]) to remove the effects of illumination conditions.

$$I_o = \frac{I_i - m}{\sigma}. \quad (10)$$

Where I_i is the input image, I_o is the output image, m is the mean of gray values across the input image, σ is the standard devation of them.

Like the Brodatz database case, two experiments are carried out on CUReT database to assess the performance of our method: 1. 552 images are used as training samples, and the rest 552 ones as testing samples. The proportion of training samples to testing samples is 2:2. 2. 276 images are used as training samples, and the rest 828 ones as testing samples. The proportion of training samples to testing samples is 1:3.

Each image is decomposed with different number of scales, from1 to 5. The experimental result is shown in Table 2.

It is obvious, the same to Brodatz database case, that the b-

TABLE I. CLASSIFICATION ACCURACY ON BRODATZ DATABASE(%)

scale		1	2	3	4	5
3:3	**RD**	73.56	96.31	**99.52**	99.52	99.52
	GD	71.79	94.07	96.15	**96.96**	95.35
1:5	**RD**	49.23	82.40	97.79	**98.46**	98.27
	GD	64.13	83.27	**94.52**	94.13	89.71

TABLE II. CLASSIFICATION ACCURACY ON CURET DATABASE(%)

scale		1	2	3	4	5
2:2	**GD**	64.86	83.33	**86.23**	85.51	80.43
1:3	**GD**	62.56	82.61	**86.11**	85.39	76.09

est performance is achieved with 3 scales of DWT decomposition. For the more complex (or more real) situation of CUReT database, still a good result is obtained with our method. The classification accuracy reached 86.23% and 86.11% corresponding to the training testing proportion of 2:2 and 1:3 respectively.

VII. Conclusion

In order to realize rotation invariant, a dominant orientaion estimation method with Gabor filters is proposed. Images are modulated with the estimated dominant orientaions, and then features are extracted with DWT. Classification experiments with nearest neighbor classifier are carried out on Brodatz database and CUReT database，and high accuracy is obtained. Small amount of operations (or high speed or low computational cost) characterizes the proposed method with practical utility and a solution to the multi-direction problem is given in this method. In future works, other directional filter banks could be used to estimate dominant orientation and the vote strategy could be reformed to achieve better performance.

References

[1] Pradeep M. Patil, Jayant V. Kulkarni, "Rotation and intensity invariant shoeprint matching using Gabor transform with application to forensic scienc," Pattern Recognition, Vol. 42, pp. 1308–1317, 2009.

[2] Ramchandra Manthalkar, P. K. Biswas, B. N. Chatterji, "Rotation invariant texture classification using even symmetric Gabor filters," Pattern Recognition Letters, Vol. 24, pp. 2061–2068, 2003.

[3] Ju Han, Kai-Kuang Ma, "Rotation-invariant and scale-invariant Gabor features for texture image retrieval," Image and Vision Computing, Vol. 25, pp. 1474–1481, 2007.

[4] Franci Lahajnar, Stanislav Kovačič, "Rotation-invariant texture classification," Pattern Recognition Letters, Vol. 24, pp. 1151–1161, 2003.

[5] S. Arivazhagan, L. Ganesan, S. Padam Priyal, "Texture classification using Gabor wavelets based rotation invariant features," Pattern Recognition Letters, Vol. 27, pp. 1976–1982, 2006.

[6] Minh N. Do, Member, IEEE, and Martin Vetterli, Fellow, IEEE, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden Markov models," IEEE Transactions on Multimedia, Vol. 4, No. 4, pp. 517–527, 2002.

[7] Kourosh Jafari-Khouzani, Hamid Soltanian-Zadeh, Senior Member, IEEE, "Radon transform orientation estimation for rotation invariant texture analysis," Technical Report TR-0105, pp. 1–20, Jan 12 2005.

[8] Peiling Cui, Junhong Li, Quan Pan, Hongcai Zhang, "Rotation and scaling invariant texture classification based on Radon transform and multiscale analysis," Pattern Recognition Letters, Vol. 27, pp. 408–413, 2006.

[9] Xuan Wang, Fang-xia Guo, Bin Xiao, Jian-feng Ma, "Rotation invariant analysis and orientation estimation method for texture classification based on Radon transform and correlation analysis," J. Vis. Commun. Image R, Vol. 21, pp. 29–32, 2010.

[10] Kourosh Jafari-Khouzani and Hamid Soltanian-Zadeh, Senior Member, IEEE, "Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms," IEEE Transactions on Image Processing, Vol. 14, No. 6, pp. 783–795, 2005.

[11] Xuan Wang, Bin Xiao, Jian-Feng Ma, Xiu-Li Bi, "Scaling and rotation invariant analysis approach to object recognition based on Radon and Fourier-Mellin transforms," Pattern Recognition, Vol. 40, pp. 3503–3508, 2007.

[12] Zhenhua Guo, Lei Zhang, David Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," Pattern Recognition, Vol. 43, pp. 706–719, 2010.

[13] Timo Ojala, Matti Pietikäinen, Senior Member, IEEE, and Topi Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp. 517–527, July 2002.

[14] Wumo Pan, T.D. Bui, C.Y. Suen, "Rotation invariant texture classification by ridgelet transform and frequency-orientation space decomposition," Signal Processing, Vol. 88, pp. 189–199, 2008.

[15] R. Manthalkar, P.K. Biswas, B.N. Chatterji, "Rotation and scale invariant texture features using discrete wavelet packet transform," Pattern Recognition Letters, Vol. 24, pp. 2455–2462, 2003.

[16] Chi-Man Pun, "Rotation-invariant texture feature for image retrieval," Computer Vision and Image Understanding, Vol. 89, pp. 24–43, 2003.

[17] Manik Varma, Andrew Zisserman, "A statistical approach to texture classification from single images," International Journal of Computer Vision, Vol. 62(1/2), pp. 61–81, 2005.